AD-A092 139   NAVAL POSTGRADUATE SCHOOL MONTEREY CA   F/G 4/2
IMPROVEMENT IN TROPICAL CYCLONE FORECASTS BY MULTIPLE LINEAR RE--ETC(U)
JUN 80   K A PETERSON

UNCLASSIFIED   NL

1 OF 1
AD A092139

END
DATE
FILMED
81-2
DTIC

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

IMPROVEMENT IN TROPICAL CYCLONE FORECASTS
BY MULTIPLE LINEAR REGRESSION EQUATION
ADJUSTMENT OF ANALOG TRACKS

by

Kenneth Allen Peterson

June 1980

48

Thesis Advisor: R.L. Elsberry

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A092	3. RECIPIENT'S CATALOG NUMBER 139
4. TITLE (and Subtitle) IMPROVEMENT IN TROPICAL CYCLONE FORECASTS BY MULTIPLE LINEAR REGRESSION EQUATION ADJUSTMENT OF ANALOG TRACKS		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis June 1980
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kenneth Allen Peterson		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93940		12. REPORT DATE June 1980
		13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Naval Postgraduate School Monterey, CA 93940		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A statistical technique is tested for adjusting the analog storm tracks used in the Northwest Pacific Analog Tropical Cyclone Forecasting Program (TYAN78). In the present version the tracks are rotated based on linear extrapolation of the deviations of the past 12-h motion for the analog and the current storm. The technique proposed here uses multiple linear regression equations to adjust the 12-h through 72-h analog positions. Predictors include translation speeds along the analog track from the -36-h to 72-h positions plus

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
(Page 1) S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

the comparison of the past 12-, 24- and 36-h translated positions of the analog with the current storm. By simply averaging the regression adjusted positions, the vector forecast errors were smaller than the TYAN78 errors. Both the TYAN78 and regression adjusted forecast errors were less than the official JTWC forecast errors. The improvement in the regression adjusted forecasts was in a smaller standard deviation relative to the mean position. For example, in the dependent sample the mean errors at 72 h for the official forecast, TYAN78 forecast and regression adjusted forecast were 438, 321, and 288 nautical miles, respectively. The corresponding standard deviations were 198, 151, 162 nautical miles.

Approved for public release; distribution unlimited

Improvement in Tropical Cyclone Forecasts By Multiple Linear Regression Equation Adjustment of Analog Tracks

by

Kenneth Allen Peterson
Captain, United States Air Force
B.A., Southern Illinois University, 1968
M.S., University of Southern California, 1977

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN METEOROLOGY

from the

NAVAL POSTGRADUATE SCHOOL
June 1980

Author ______________________

Approved by: ______________________
Thesis Advisor

____
Chairman, Department of Meteorology

____
Dean of Science and Engineering

ABSTRACT

A statistical technique is tested for adjusting the analog storm tracks used in the Northwest Pacific Analog Tropical Cyclone Forecasting Program (TYAN78). In the present version the tracks are rotated based on linear extrapolation of the deviations of the past 12-h motion for the analog and the current storm. The technique proposed here uses multiple linear regression equations to adjust the 12-h through 72-h analog positions. Predictors include translation speeds along the analog track from the -36-h to 72-h positions plus the comparison of the past 12-, 24- and 36-h translated positions of the analog with the current storm. By simply averaging the regression adjusted positions, the vector forecast errors were smaller than the TYAN78 errors. Both the TYAN78 and regression adjusted forecast errors were less than the official JTWC forecast errors. The improvement in the regression adjusted forecasts was in a smaller standard deviation relative to the mean position. For example, in the dependent sample the mean errors at 72 h for the official forecast, TYAN78 forecast and regression adjusted forecast were 438, 321, and 288 nautical miles, respectively. The corresponding standard deviations were 198, 151, 162 nautical miles.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author wishes to thank his wife, Nancy, for her patience throughout the course of this research and her help in organizing the data to be used in the final stages, and for helping to type the draft.

The author thanks Professor R.L. Elsberry for his patience and help during this research.

Many thanks go to LT Jim Bradshaw (USN) for his explanations of the use of the CDC 6500 and interactive terminal usage. With his help the work effort was greatly reduced.

Special thanks to Michael Beck for his help in setting up the TRS-80 as an interactive terminal which was used heavily both on the Fleet Numerical Oceanography Center's (FNOC) development computer (HAL) and on the W.R. Church Computer Center time sharing system.

The author wishes to thank Messrs. Bob Alden and Warren S. Yogi for their assistance and loan of Ocean Data Systems, Inc., documentation of the Analog Tropical Cyclone Forecasting Program (TYAN78).

The financial support of the Naval Air Systems Command and computer support by Fleet Numerical Oceanography Center and W.R. Church Computer Center are acknowledged.

I. INTRODUCTION

Analog tropical cyclone prediction models have been in operational use since 1970 (Jarrell and Wagoner, 1973). The basic principle of analog models is to use the mean position of all historical storms which have characteristics similar to the current storm. Probability ellipses are generated as part of the output forecast based on means and standard deviations of the latitudes and longitudes and the correlation coefficient between them (Jarrell and Wagoner, 1973).

Fleet Weather Central/Joint Typhoon Warning Center in Guam (FWC/JTWC) has used an analog model named TYFOON since 1970. TYFOON used a weighted mean latitude and longitude for the predicted 24-, 48-, and 72-h positions. A modified version of TYFOON named TYFOON-72 was instituted in 1972. The important modifications were improved input/output efficiency, derivation of adjustment coefficients and an increased number of candidate analogs. The historical data base for TYFOON and TYFOON-72 was the period 1949-1969. In 1975 a further modification of the analog model (TYFN-75) was introduced that incorporated a data base of storm tracks for the Northwest Pacific for the years 1945 to 1973. The most recent update and modification to the Analog Tropical Cyclone Forecast Program is named TYAN78. This version is the result of conversion from the Control Data Corporation (CDC) 3100 computer system to the

CDC 6500 system (Fleet Numerical Oceanography Center 6500/ Naval Environmental Display Station-1). In this version, as in the earlier ones, the analog tracks are translated so that the best position of the analog is co-located with the current position. The remaining analog positions are adjusted accordingly and the track is then rotated based on linear extrapolation of the deviations of the past 12-h motion for the analog and the current storm. An updated data base is used in TYAN78 that includes storm tracks for the Northwest Pacific from 1945 through 1976. New weighting coefficiencts are used based on these data.

The purpose of this thesis was to derive multiple regression equations, based on analogs selected by TYAN78 for the Northwest Pacific, to adjust the 12-h through 72-h analog positions. The objective of using regression-adjusted positions is to decrease the dispersion about the actual position.

In addition to simply averaging the regression-adjusted positions, a weighting scheme based on standard deviations of the vector lengths found in comparing the 12-, 24-, and 36-h past positions of the analogs to the current storm was tested. The results showed no improvement over the averaged regression-adjusted analog forecast for the 24- and 48-h forecasts. There was a very slight improvement in the 72-h forecast that indicates a step toward possibly using recent track information to select the best analogs for use in TYAN78.

II. MODEL AND SAMPLE

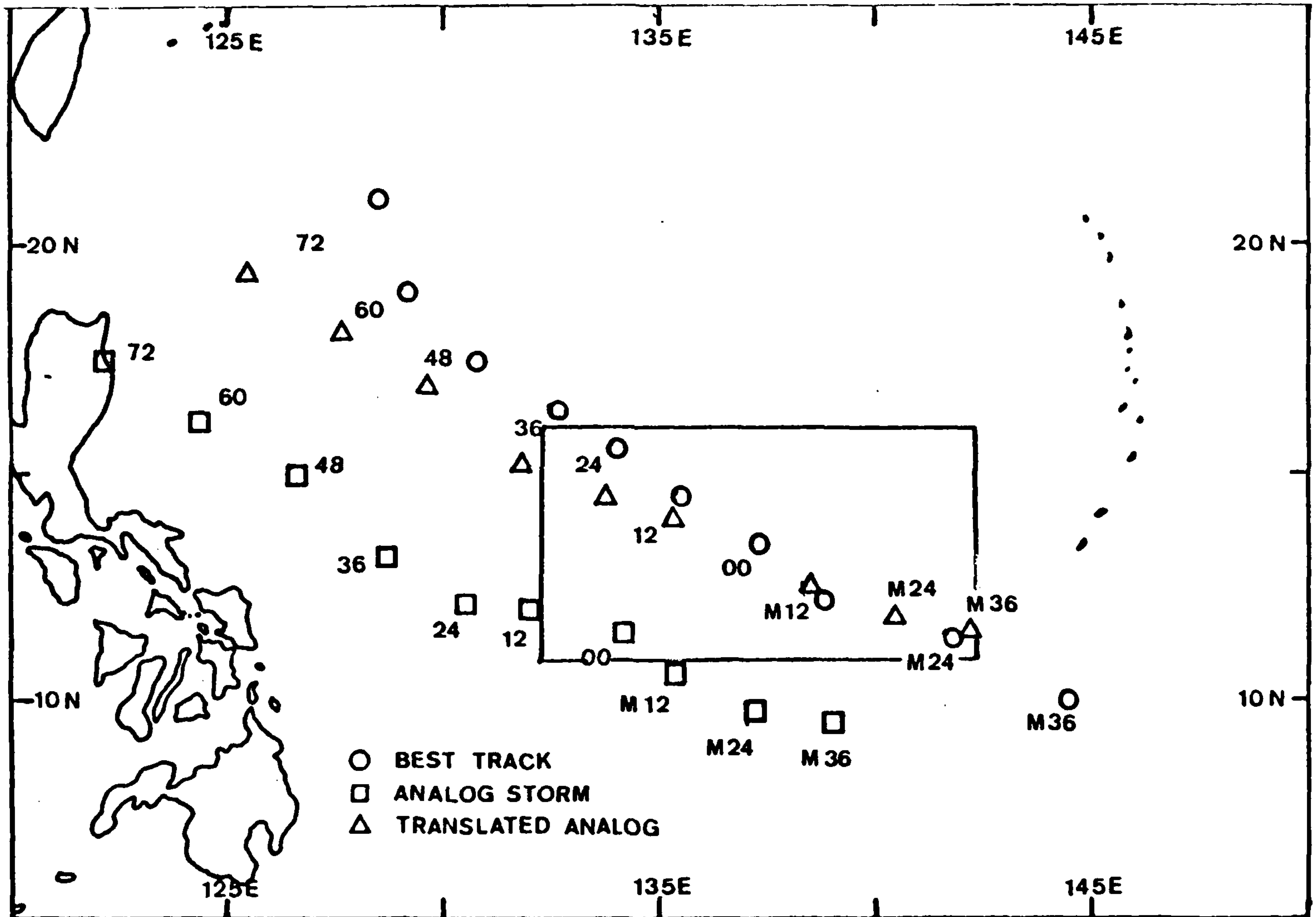
A. THE TYAN78 MODEL

The model used to generate the data for this research is the TYAN78 model resident on the FNOC CDC 6500 computer. The TYAN78 program is a result of a conversion and combination of the five analog tropical cyclone models used by FWC/JTWC, Guam (ODSI, 1978). The program conversion was from the CDC 3100 system to the CDC 6500 system. In addition to the conversion, five tropical cyclone analog models were combined to form TYAN78. The five were analog forecast models for the Northwest Pacific, Northeast Pacific, Southwest Pacific, Southwest Indian and North Indian Ocean regions.

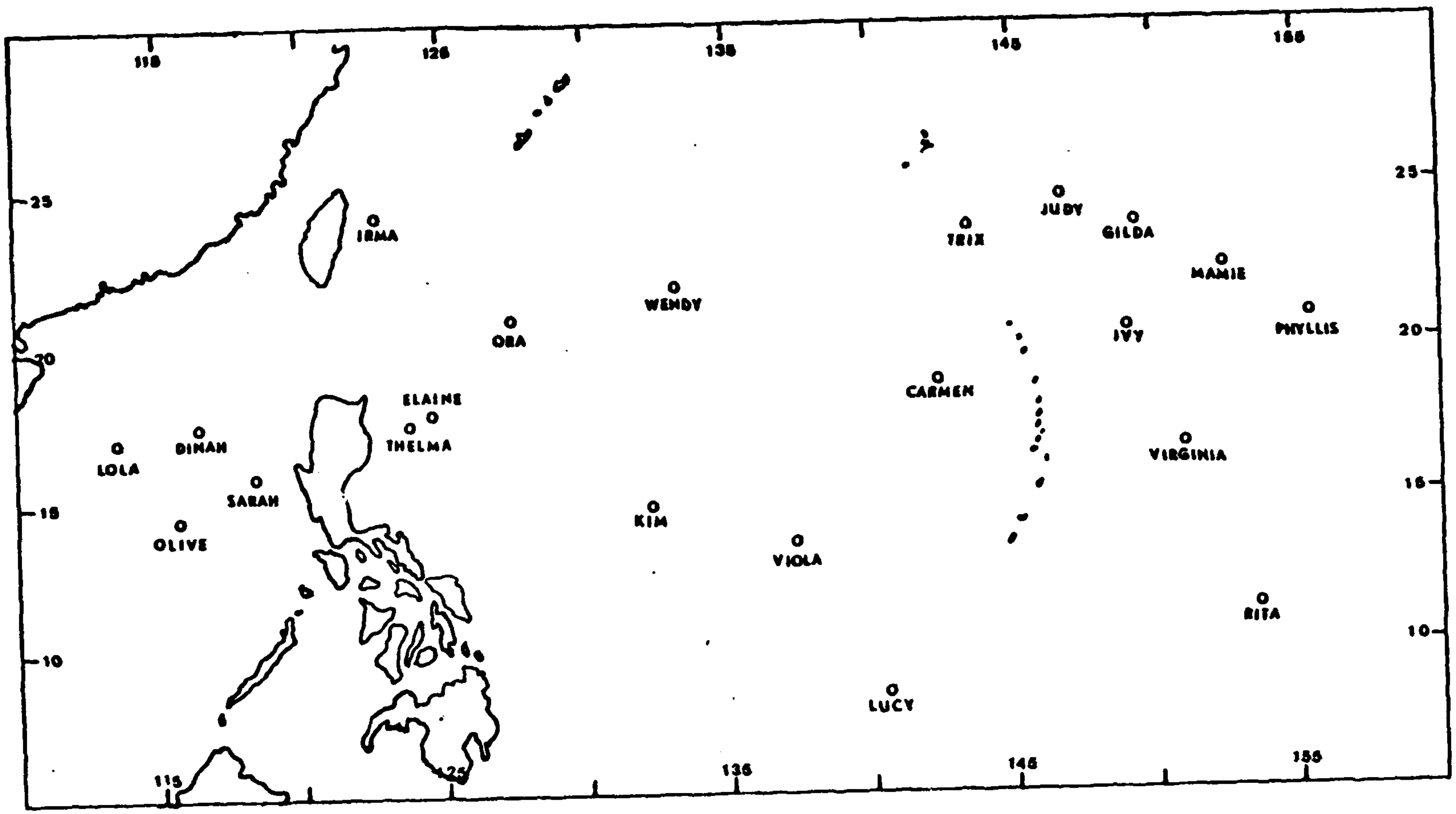
The TYAN78 program was designed to operate without computer operator intervention. The forecaster enters the required input information through Automatic Response to Query (ARQ) and receives a forecast for tropical cyclones in the aforementioned ocean basins. For the Northwest Pacific basin, the forecaster has the option of selecting only recurver analogs, only non-recurver analogs or a combination of all potential analogs. Separate TYAN78 forecasts are produced for each case. In the present work the combined sample was chosen.

The TYAN78 program selects potential analogs as described by Jarrell and Wagoner (1973), except that the time envelope

is ± 35 days rather than ± 50 days. Basically, the model searches the history file for all storms that occurred within the time envelope relative to the current (input) storm. Then the model checks for analog positions that occurred within ± 2.5 degrees latitude and ± 5.0 degrees longitude, as illustrated in figure 1. The time and space envelope will include more than one position along the storm track. A similarity index is calculated for each position within the envelope. The similarity index is defined as $SI = \Sigma\, W_i \cdot |P_{ic} - P_{ia}|$, where W_i is the weight factor associated with the ith parameter, P_{ic} is the value of the ith parameter of the current storm and P_{ia} is the value of the ith parameter of the analog position. The W_i values were originally specified using "educated guesses" for each parameter (Jarrell and Wagoner, 1973). The parameters used are listed in Table I. The "best" position for each analog storm is defined as the one with the lowest simliarity index (Jarrell and Wagoner, 1973). The latitude and longitude correction necessary to translate the "best" position of the analog to the current position of the input storm is then calculated. This vector displacement is then added to all the points along the corresponding analog track. Thus the "best" positions of all analogs are initially co-located with the current storm position. The TYAN78 program then ranks the potential analog storms with the highest rank given to the storm with the lowest similarity index. This ranking is used in finding a weighted mean position for each set of forecasts.



1. Illustration of best track and analog before and after translation of best position of analog to current position of input storm (Viola 1800 GMT 19 November 1978).



2. Map of the dependent sample input "current" positions.

TABLE 1

Parameters for each 6-h position of the tropical storms and typhoons from which similarity indices are calculated (Jarrell and Wagoner, 1973).

Parameter	Units
Julian date	
Latitude	nearest tenth of degree
Longitude	nearest tenth of degree
12 h direction of movement	nearest degree
12 h speed of movement	nearest knot
18 h direction of movement	nearest degree
18 h speed of movement	nearest knot
24 h direction of movement	nearest degree
24 h speed of movement	nearest knot
48 h direction of movement	nearest degree
48 h speed of movement	nearest knot
Radius of outer closed isobar	nearest degree
12 h change of radius	nearest degree
Sea-level pressure (minimum)	nearest millibar
12 h change of min. SLP	nearest millibar
Maximum wind speed	nearest 5 knots
Minimum 700-mb height	nearest decameter
Latitude of 700-mb ridge to north	nearest degree
Height of 700-mb ridge to north	nearest decameter
Longitude of 700-mb trough to the west at 35N	nearest degree
Height of 700-mb trough at 35N	nearest decameter

In this research the TYAN78 model was used to the point of selecting and ranking the analogs. The normal TYAN78 forecasts were also collected for later comparison with the new methods tested.

The Program Maintenance Manual (ODSI, 1978) provided a means for having some of the coefficients and similarity indices printed out. A list of the top five analogs with the best position identified was also available as an option. (Note that this option is now part of the output to the forecaster after a recent modification by ODSI to TYAN78). The TYAN78 program was modified to list all analogs that were included in the search, and to print out the top 20 analogs including the "best" position for each storm. From this list of potential analogs, the positions for all analogs which included the -12-, -24-, -36-h positions and the +12-, 24-, 36-,28-,60-,72-h positions were retained.

B. SELECTION OF INPUT STORMS

The data base for the Northwest Pacific includes storms from 1945 through 1976. It is thus possible that use of a storm during this period would result in its selection as a potential analog. A sample storm from 1968 (NINA 1200GMT 22 November as initial position) was used in the ARQ input format to generate the forecast and print out the top five analogs. The option of having all three forecast modes was entered in the ARQ. In the straight option, the input storm was selected

as an analog and was ranked number two rather than one, and the "best" position was the current (input) position. However, in the option that considered all storms, the input storm was ranked number five, and the "best" position was 18 h after the input position. Because of these results, it was decided not to use any sample storms that occurred during the data base period (1945-1976). Thus, the dependent sample consists of storms from 1977 and 1978.

From the 1977 and 1978 data, 21 storms were selected to give a representative geographic distribution (see figure 2). For each of these storms, the name, date-time group, initial position and 12-h past position (all minimum ARQ requirements) were entered. Since all of the sample storms had at least a 48-h history, the 18-, 24-, and 48-h past positions were also included. All position information was best track information from the Annual Typhoon Reports (1977 and 1978). The maximum surface wind at the initial time was used with initial latitude to derive a minimum sea level pressure, and a 700-mb height from figure 9-15 in Atkinson (1971). Other parameters such as the latitutde of the 700-mb ridge (see Table 1) were not available as input to the analog selection method.

Using the above parameters as the TYAN78 ARQ input, 21 storms were processed on the FNOC computer. Appendix A describes the procedures used for processing the data. The output thus included 21 forecasts for 24, 48, and 72 h. The listing of the 20 best analogs for each

storm included a total of 308 analogs with at least a 36-h history. Only 308 analogs were available because: 1) not all input storms had 20 potential analogs; and 2) not all of the potential analogs had the 36-h history required for application of the statistical method. This sample of 308 analogs and the associated best tracks of the 21 input storms was used to develop the multiple linear regression equations.

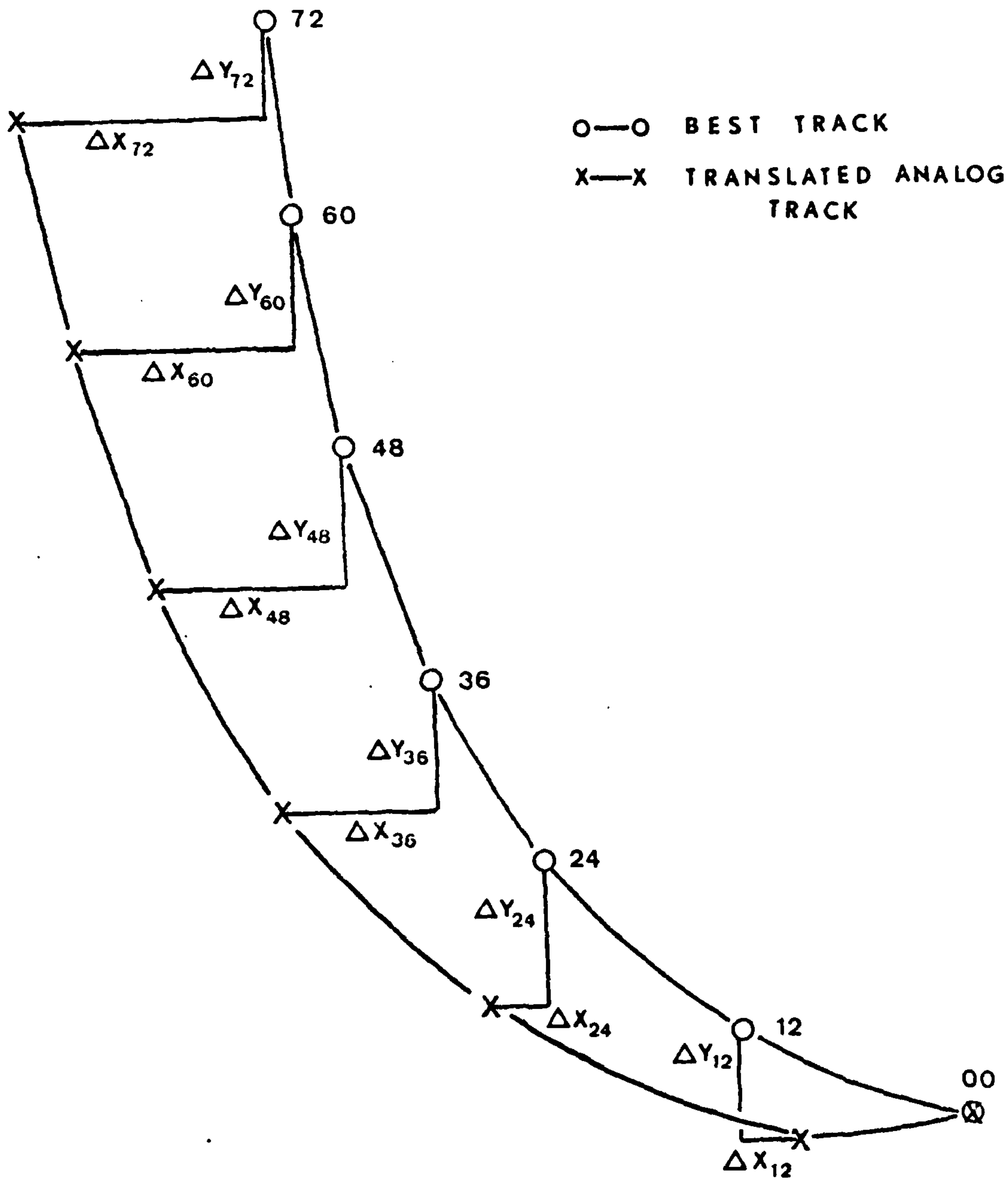
III. DEVELOPMENT OF MULTIPLE LINEAR REGRESSION EQUATIONS

As described above, each of the potential analogs was translated so that the "best" position was co-located with the initial position of the corresponding input storm. Consider an analog storm track and the best track of the input storm as in figure 3. The best track positions are considered known for the dependent sample used to develop the regression equations, and represent the desired forecast positions after adjustment of the analog positions. Thus the predictands are the zonal and meridional displacements to be added to the analog positions. A total of 12 regression equations would normally comprise these sets. However, a provision is made for alternative regression equations if the analog did not persist for 72 h. Thus there is one set each for when the analog storm persisted 72, 60, 48, 36, or 24 h past the best position.

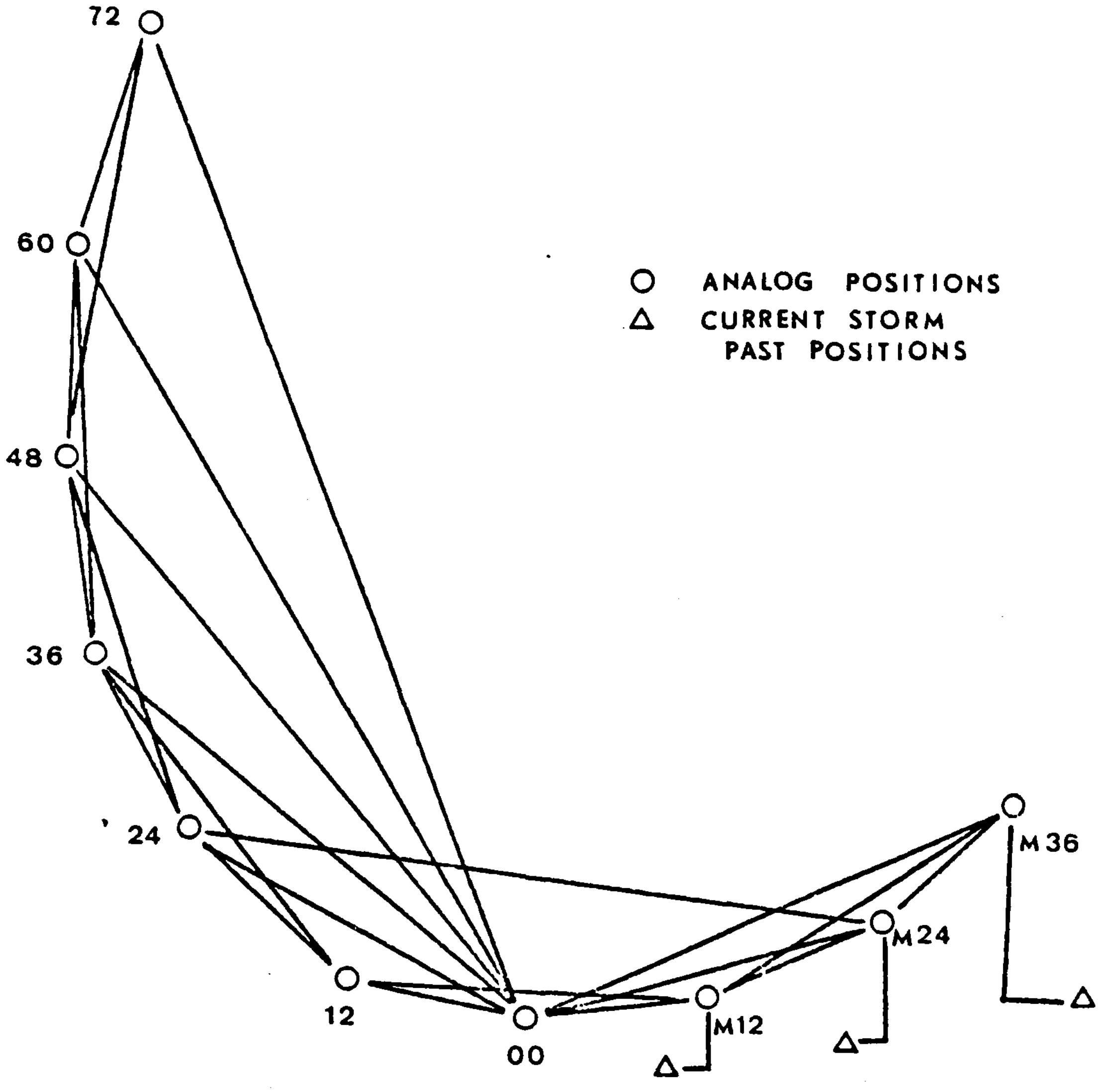
The predictors used in this research are based on the work by Elsberry and Frill (1979). They were able to improve the Tropical Cyclone Model (TCM) forecasts when the tracks were adjusted with regression equations that included past position information. A similar approach is used in this study, which required that only analogs with a 36-h history be retained in the sample. A schematic of the predictors, with line segments representing displacement vectors, is shown in figure 4. Table 2 lists the predictors and predictands. These vectors were

converted to zonal and meridional velocity components. Zonal and meridional velocity predictors were also computed for -12-, -24-, and -36-h positions of the analog compared to the input storm's corresponding past positions. Thus there are 52 potential predictors.

The regression equations were generated by the Stepwise Regression Program of the Biomedical Computer Programs package (Dixon and Brown, 1977). The F-to-enter and F-to-remove values were specified as 4.0 and 3.9, respectively. These are the limits at which stepping is terminated. That is, no more variables are entered or removed if the calculated F-to-enter or F-to-remove fall below 4.0 or 3.9, respectively. In the final choice of equations, a maximum of five predictors was allowed, although no predictor was included if it contributed less than .01 to the increase in the multiple correlation coefficient squared (r^2). The equations are presented in Table 3 , together with the multiple correlation coefficient squared, the F-ratio and degrees of freedom, and the standard error of estimate. The F-ratio used here is a test statistic defined as the ratio of the mean square due to regression to the mean square due to error. It is available from the analysis of variance table provided by the stepwise regression program. Under the null hypothesis this statistic has an F distribution with numerator degrees of freedom equal to the number of variables selected thus far. The denominator degrees of freedom equals the sample size minus the number of variables



3. Schematic of the predictands (ΔX_{12}, ΔY_{12}, ΔX_{24}, ΔY_{24}, etc.) which are the differences between best track and analog positions after translation.



4. Illustration of the intervals over which the predictors were computed. Positions are shown at 12-h intervals.

TABLE 2

Predictors/predictands used to develop regression equations for analog track adjustment

1. Predictands: ΔX, ΔY (nautical miles)

   Times at which predictands are computed:

   12, 24, 36, 48, 60, 72 hrs

2. Predictors: u, v (knots)

   Time intervals over which each predictor was computed:

   00-12, 00-24, 00-36, 00-48, 00-60, 00-72, 12-24, 12-36, 24-36, 24-48, 36-48, 36-60, 48-60, 48-72, 60-72

   00-M12, 00-M24, 00-M36, M12-M24, M12-M36, M12-12, M24-24, M24-M36

   where 00 = best (current) position

   M24 = 24 hours before current position

   12 = 12 hours from current position

   M12-B12, M24-B24, M36-B36

   where B24= 24 h past position of current storm.

selected minus one. Further details of the F distribution may be found in any statistics text. The null hypothesis that the set of coefficients for each equation is zero can be rejected for the F statistics and degrees of freedom shown. That is, the variables may be regarded as significant in predicting their respective predictands at a significance level of .01.

One should note that predictors based on past positions were included in each of the equations. Note especially the occurrence of predictors involving the -12-h and current (00) position of the analog. The predictors that compare the -12-h position of the analog with the -12-h position of the current storm are also chosen frequently. Thus the regression equation approach developed here uses more than just the past 12-h motion to adjust the analog tracks.

The table of equations only lists those regression sets for analog storms that existed either 72 h, 48 h, or 24 h past the time of the best position. Equations for intermediate times such as 60 h and 36 h were also developed and are evaluated below, but they are not listed here. The tables include the multiple correlation coefficient squared (r^2), F-ratio with numerator and denominator degrees of freedom, and standard error of estimate at the step the equation was developed. For each of the equations the computed F-ratio and r^2 are consistently less for the zonal displacement prediction than for the meridional displacement. This indicates that the linear

TABLE 3a

Regression equations for adjusting the zonal and meridional displacements (n.mi.) of the analog storms, multiple correlation coefficient squared (r^2), F-ratio with numerator and denominator degrees of freedom, and the standard error of estimate at the step the predictors were selected. These equations were produced by BMDP2R (Stepwise Regression Program) [Dixon and Brown, 1977]. Parameters are listed in order of selection.

a. Equation set for analog storms that persist to 72 h. These are based on 21 storms that selected 177 analogs which existed 72 h past the "best" position.

MULTIPLE LINEAR REGRESSION EQUATIONS	MULTIPLE CORRELATION COEFFICIENT SQUARED (r^2)	F-RATIO (DEGREES OF FREEDOM)	STANDARD ERROR OF ESTIMATE
DXER72 = 147.435 + 34.502 (VXM200) - 2.171 (BXER12) - 51.198 (VX0072) - 41.851 (VYM6M4) + 1.134 (BYER36)	.54	39.9 (5,171)	324.9
DYER72 = 127.735 - 64.768 (VY0072) - 3.146 (BYER12) + 49.109 (VYM200) + 1.729 (BYER36) - 2.681 (BYER24)	.78	123.8 (5,171)	166.1
DXER48 = 43.576 - 50.299 (VX2448) - 1.635 (BXER12) + 16.064 (VXM200) + 21.487 (VX3660) + 7.37 (VY0012)	.59	48.8 (5,171)	168.7
DYER48 = 86.833 - 2.506 (BYER12) + 31.469 (VYM200) - 45.67 (VY0048 + .813 (BYER36) - .964 (BYER24)	.80	137.7 (5,171)	99.6
DXER24 = 51.839 - 1.972 (BXER12) + 35.632 (VXM200) - 42.64 (VXM424) + .598 (BXER24) - .28 (BYER12)	.59	48.6 (5,171)	76.9
DYER24 = 69.323 - 1.789 (BYER12) + 21.673 (VYM200) - 24.335 (VY0024) + .412 (BYER36) - 11.117 (VYM600)	.81	141.9 (5,171)	46.8

TABLE 3b

Same as 3a except for analog storms that persisted to 48 h past "best" position. These are based on 21 storms with 242 analogs.

MULTIPLE LINEAR REGRESSION EQUATIONS	MULTIPLE CORRELATION COEFFICIENT SQUARED (r^2)	F-RATIO (DEGREES OF FREE-DOM)	STANDARD ERROR OF ESTIMATE
DXER48 = 50.519 - 16.747 (VX2448) - 2.053 (BXER12) + 21.173 (VXM200) - 19.392 (VX1224)	.73	159.7 (4,237)	161.9
DYER48 = 91.881 - 45.062 (VY0048) - 3.213 (BYER12) + 31.272 (VYM200) + .324 (BYER36)	.81	246.9 (4,237)	102.5
DXER24 = 49.966 - 2.031 (BXER12) + 31.605 (VXM200) - 22.05 (VX0024) + .366 (BXER36) - 16.138 (VXM600)	.69	103.6 (5,236)	73.2
DYER24 = 79.702 - 23.953 (VY0024) - 1.707 (BYER12) + 21.107 (VYM200) + .438 (BYER36) - 13.432 (VYM600)	.81	197.3 (5,236)	48.2

TABLE 3c

Same as 3a except for analog storms that persisted to 24 h past "best" position. These are based on 21 storms with 308 analogs.

MULTIPLE LINEAR REGRESSION EQUATIONS	MULTIPLE CORRELATION COEFFICIENT SQUARED (r^2)	F-RATIO (DEGREES OF FREEDOM)	STANDARD ERROR OF ESTIMATE
DXER24 = 11.375 - 11.355 (VX1224) - 1.708 (BXER12) + 26.395 (VXM200) - 18.21 (VXM212) + .144 (BXER36)	.68	126.9 (5,302)	75.7
DYER24 = 76.272 - 23.305 (VY0024) - 1.746 (BYER12) + 19.827 (VYM200) + .452 (BYER36) - 12.261 (VYM600)	.81	254.9 (5,302)	48.0

dependence of the meriodional displacements on the predictors is stronger than that of the zonal displacements.

To make the regression-adjusted forecasts, the zonal and meriodional displacements calculated from the multiple linear regression equation were used to correct the analog storm's 24-, 48- and 72-h positions. These adjusted positions were then averaged to provide a forecast position. Note that these forecasts are based on 20 or fewer analog storms and that the analogs must have had at least a 36-h history before the "best" position. This is not the same number of analogs as TYAN78 would evaluate, and these forecasts are not weighted as in TYAN78.

A weighting scheme based on the standard deviations of the vector lengths between the current storm and analog storm -12-, -24-, and -36-h positions was also tested. The standard deviations were obtained from the sample of 308 analog storms. One weight was assigned to all corrections for zonal and meridional displacements. The equation is as follows:

$$WGT = \frac{S_{-12} + S_{-24} + S_{-36}}{(BX12^2 + BY12^2)^{1/2} + (BX24^2 + BY24^2)^{1/2} + (BX36^2 + BY36^2)^{1/2}}$$

where S_{-12} is the sample standard deviation of the vector distance between analog storm and current storm -12-h positions. Here $(BX12^2 + BY12^2)^{1/2}$ is the calculated vector distance for the -12-h position comparisons for the individual analog storms.

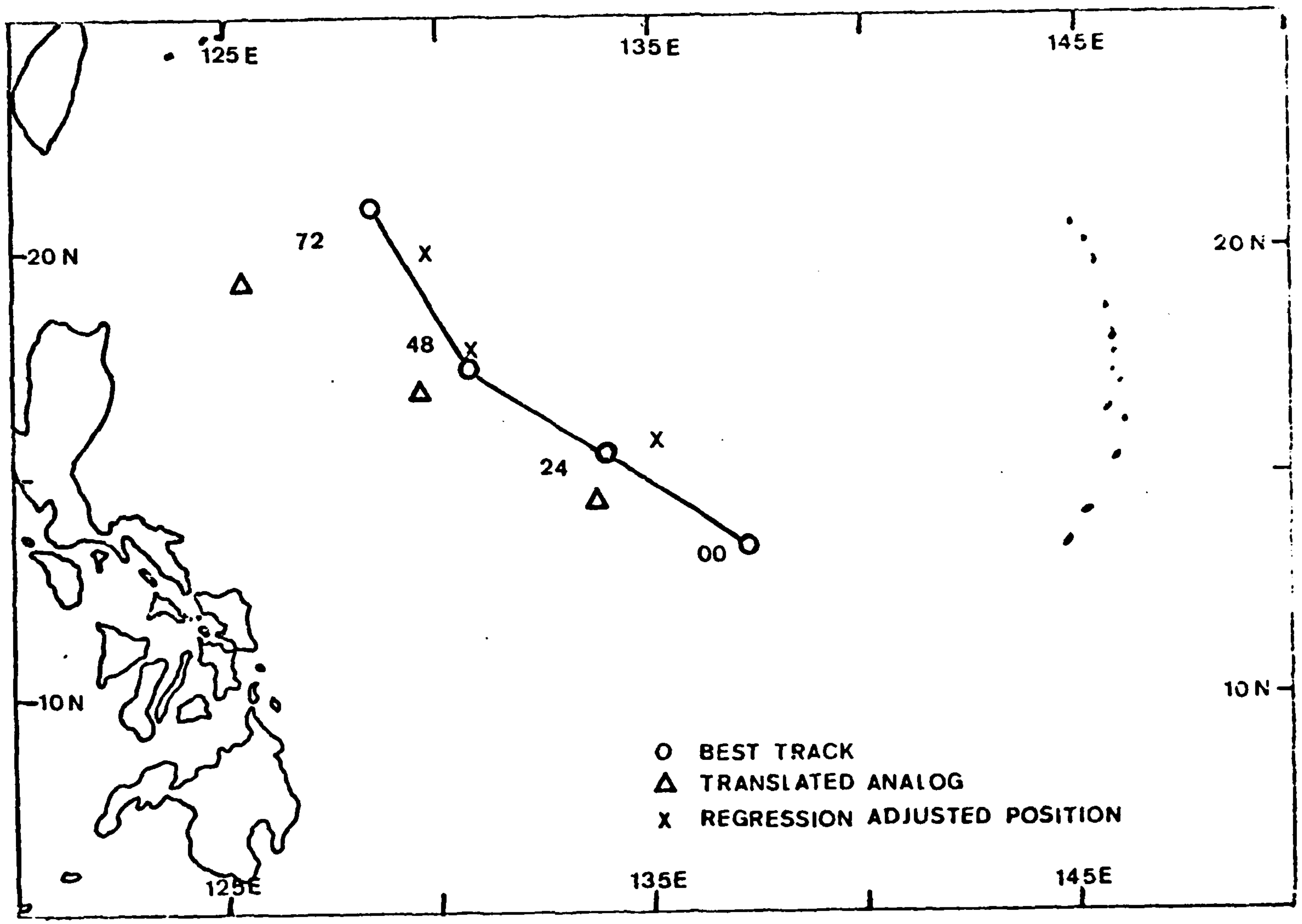
IV. RESULTS

A. DEPENDENT SAMPLE TEST

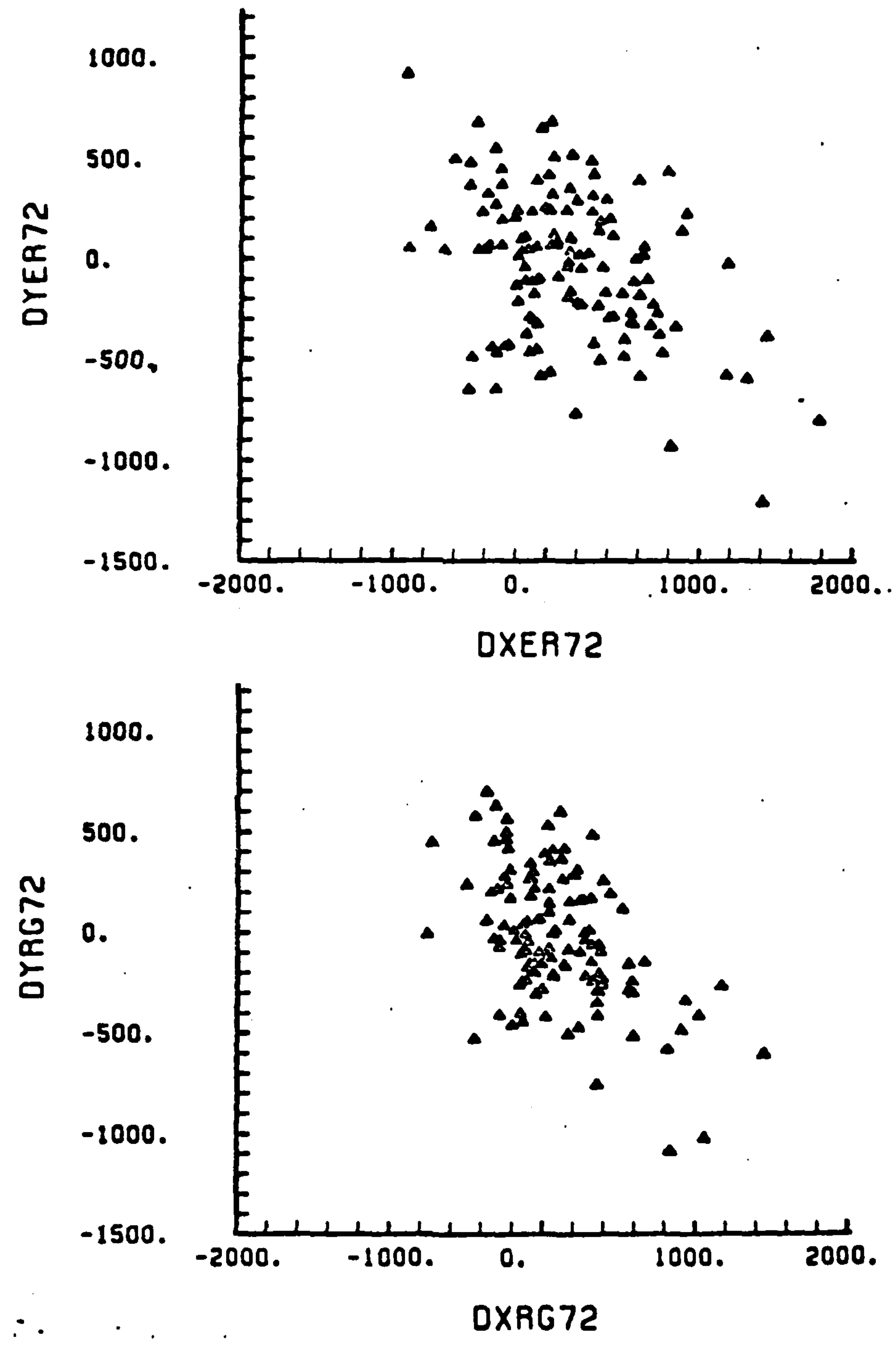
The multiple linear regression equations were initially tested using the dependent sample. Included in the dependent sample of 21 storms were two that did not have official forecast error statistics. These two storms were deleted from the evaluation below. A list of storms used is presented in Table 7.

Figure 5 illustrates an example of the analog storm positions before and after the regression adjustments compared to the best track of the input storm. The analog storm has been translated so that its "best" position is co-located with the current position of the input storm. The regression adjustment overcorrects somewhat for this example, but it is expected that the adjustments for the other analog storms will give a distribution about the actual position.

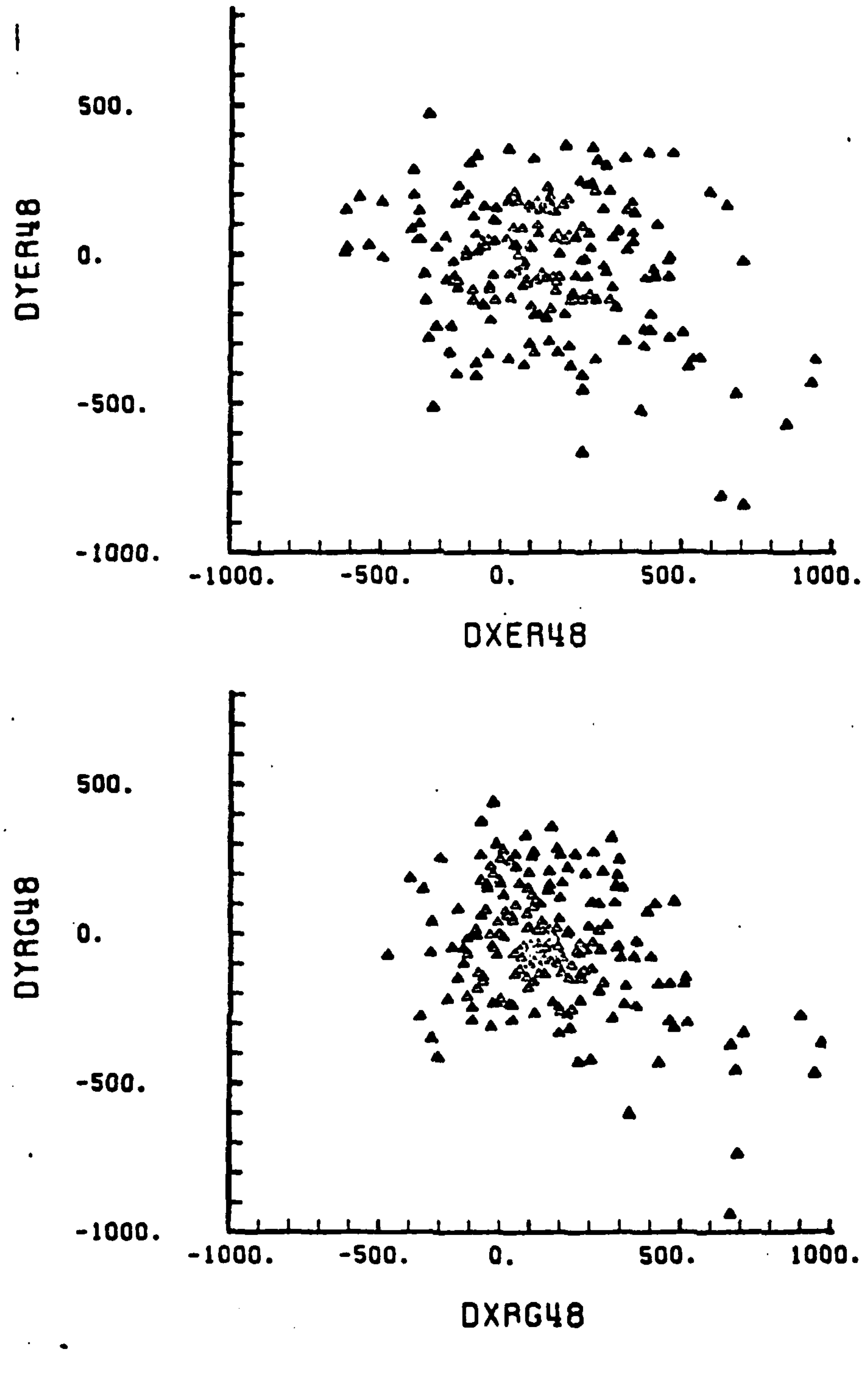
In the first test using dependent sample data, scatter diagrams were prepared showing positions of analogs relative to the actual 24-, 28-, 72-h locations (see figures 6,7, and 8). This was done for positions before and after regression to indicate the advantage of the regression adjustment over simply translating the analog. There is a reduction in the dispersion of the regressed locations, especially at 48 h. This is also demonstrated in Table 4, where the sample standard deviations of the predictands are shown to be reduced after regression.



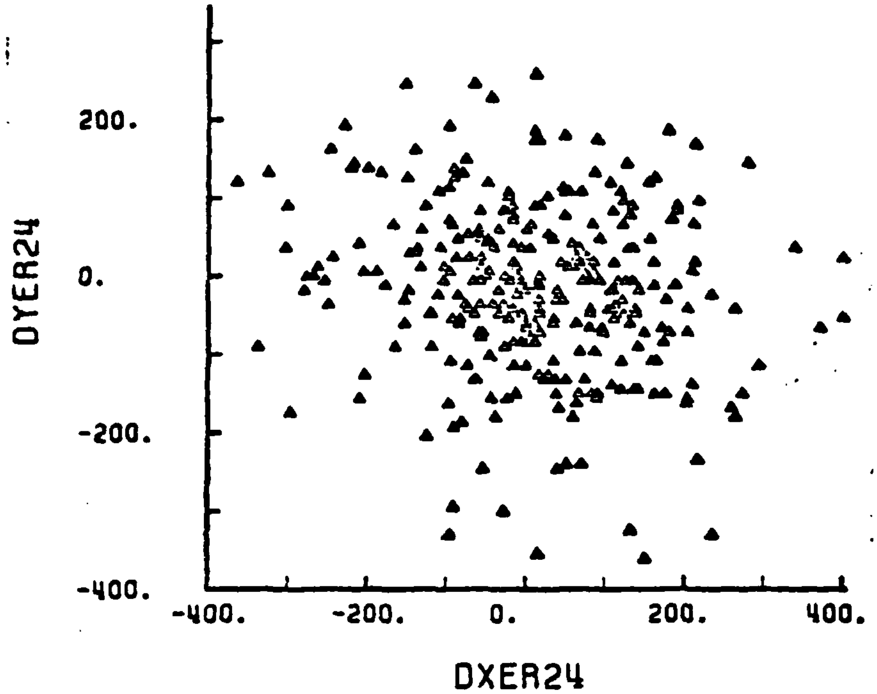
5. Illustration of an analog before regression adjustment (Δ) and after regression adjustment (X) relative to best track of current storm (O) (Viola 1800 GMT 19 November 1978).

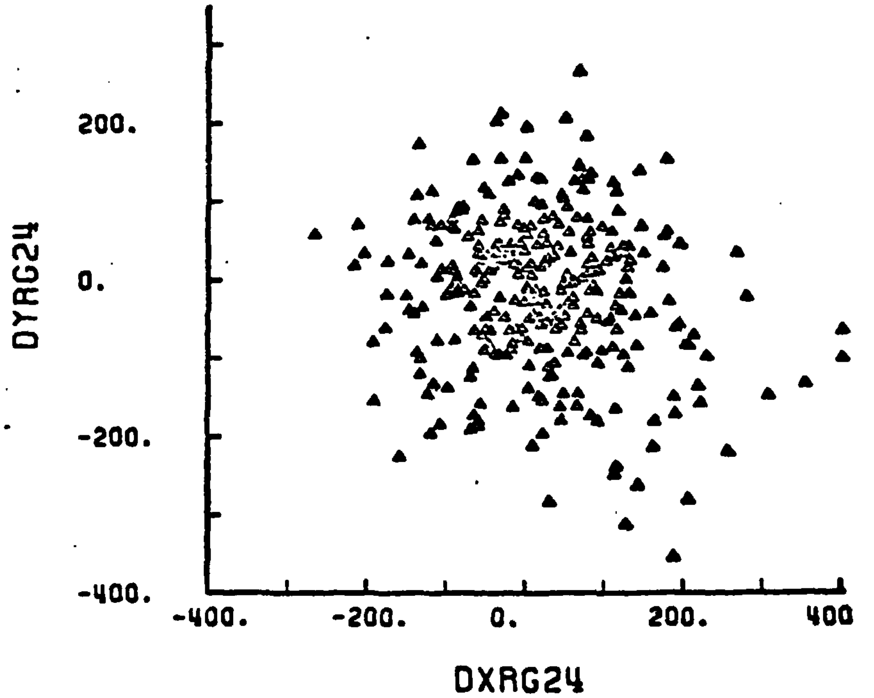


6. Scatter plot diagram of the predictands before regression (DXER72, DYER72) and after regression adjustment of the analogs (DXRG72, DYRG72) for the 72-h positions.



7. Same as figure 6 except for the 48-h positions.





8. Same as figure 6 except for the 24-h positions.

TABLE 4

Sample standard deviations of ΔX, ΔY before and after regression adjustment for 24, 48, and 72 h (nautical miles).

Predictand	Standard Deviation Before Regression	After Regression
ΔX_{24}	132.2	103.4
ΔY_{24}	108.8	98.9
ΔX_{48}	308.8	262.8
ΔY_{48}	230.9	209.9
ΔX_{72}	471.4	345.8
ΔY_{72}	351.9	311.5

In a second test, regression-adjusted forecasts and weighted regression-adjusted forecasts were made. These forecasts and the TYAN78 forecasts were compared to the corresponding best track positions and vector errors computed. The means and standard deviations of these errors are compared with those of the official forecasts in Table 5. The TYAN78 vector forecast errors at all time intervals are smaller than the official forecast errors for this sample. The regression-adjusted vector forecast errors are in turn smaller than those for TYAN78. The standard deviation for the regression-adjusted forecast errors is likewise smaller except at 72 h. Though this standard deviation for the regression-adjusted forecast error is larger than that for TYAN78, the 72-h position is better with respect to the mean position.

B. INDEPENDENT SAMPLE TEST

The ultimate test of the multiple linear regression equations is the comparison of the forecast errors for an independent data set. For the independent test a sample from 1979 was used. As in the selection of the dependent set, randomness was desired, but was not guaranteed. The same selection criteria was followed for the independent sample as for the dependent sample. The input storm was required to have at least a 108-h history and official forecast error statistics for 24, 48, and 72 h from the input initial position. A total of 18 storms were selected as the independent sample. Three were dropped from the original 18. One was dropped because the ARQ was

TABLE 5

Vector length error statistics (nautical miles) comparing official forecast errors, TYAN78 errors and regression adjusted forecast errors for the dependent sample.

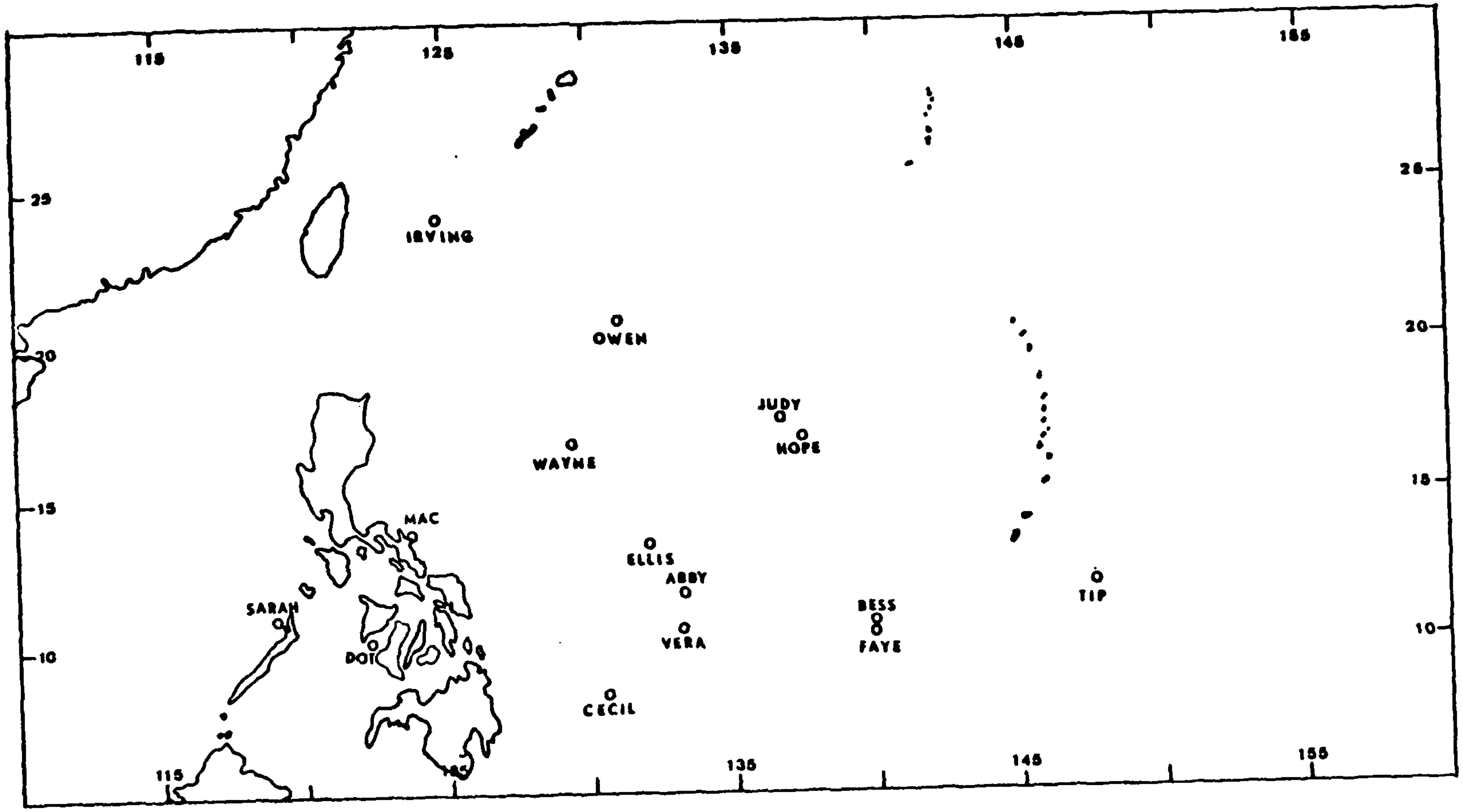
	MEAN (NM)	STANDARD DEVIATION (NM)
24 h FORECAST		
OFFICIAL	119.8	70.3
TYAN78	77.7	53.8
AVERAGED REGRESSION	68.1	42.8
WEIGHTED REGRESSION	68.8	43.8
48 h FORECAST		
OFFICIAL	238.8	110.3
TYAN78	160.2	88.5
AVERAGED REGRESSION	147.0	83.3
WEIGHTED REGRESSION	148.9	82.9
72 h FORECAST		
OFFICIAL	438.1	198.3
TYAN78	320.8	151.4
AVERAGED REGRESSION	288.0	162.1
WEIGHTED REGRESSION	287.3	161.0

rejected. Another was dropped because none of the top 20 analogs existed for 72 h past the "best" position. The third was dropped because only one analog was chosen and it did not exist for 72 h past the "best" position.

The final set of independent storms and their potential analogs was 15 storms with 199 analog storms (see figure 9 and list in Table 7). These storms and the analog storms were processed as in the dependent sample test. The resulting vector error statistics are compared in Table 6.

The means and standard deviations for the regression-adjusted forecast errors, as well as the TYAN78 forecast errors, were less than those for the official forecast errors for all time intervals. Although the mean errors for the regression-adjusted forecasts and TYAN78 forecasts were equivalent at 24 h, the stnadard deviation was reduced in the regression-adjusted forecasts. For the 48- and 72-h regression-adjusted forecasts, the standard deviations of the errors are greater than those for the TYAN78 forecasts. However, the 48- and 72-h regression-adjusted positions are better relative to the mean position.

It should be noted that the TYAN78 uses a larger sample of analog storms and has a weighting function based on the similarity index. Thus a direct comparison of the standard deviations is not possible. To get such a comparison, the present rotation scheme in TYAN78 would have to be replaced by the regression-adjustment scheme and then proceed with the remainder of the TYAN78 forecast weighting scheme. This



9. Map of the independent sample input "current" positions.

TABLE 6

Vector length error statistics (as in Table 5) except for the independent sample.

	MEAN (NM)	STANDARD DEVIATION (NM)
24 h FORECAST		
OFFICIAL	116.6	58.6
TYAN78	83.2	48.6
AVERAGED REGRESSION	84.9	42.9
WEIGHTED REGRESSION	84.3	42.4
48 h FORECAST		
OFFICIAL	205.7	141.7
TYAN78	190.9	68.3
AVERAGED REGRESSION	155.7	93.2
WEIGHTED REGRESSION	156.4	93.9
72 h FORECAST		
OFFICIAL	341.7	278.4
TYAN78	337.6	158.9
AVERAGED REGRESSION	307.7	172.2
WEIGHTED REGRESSION	312.5	169.7

TABLE 7

List of dependent and independent sample storms used in the research.

DEPENDENT SAMPLE

	NAME	DATE-TIME-GROUP (YYMMDDHH GMT)	INITIAL POSITION (DEGREES LAT/LON)		MAXIMUM SURFACE WIND (KTS)	MINIMUM SLP (MB)
	SARAH	77071812	15.9N	118.3E	50	986
	THELMA	77072300	17.5	123.8	70	971
	DINAH	77091718	17.3	116.3	55	981
	GILDA	77100606	23.8	149.2	60	975
	IVY	77102318	20.4	149.0	65	971
	KIM	77111018	14.8	132.2	125	923
	LUCY	77120106	8.4	140.3	35	998
	OLIVE	78042206	14.3	115.6	60	979
	TRIX	78071400	23.6	143.3	35	993
	VIRGINIA	78072400	16.6	151.0	45	990
	WENDY	78072612	21.9	133.1	70	969
	CARMEN	78081218	18.8	142.4	55	982
*	ELAINE	78082218	17.9	124.6	25	1005
*	IRMA	78091118	24.1	122.7	30	996
	JUDY	78091300	24.6	146.7	40	990
	LOLA	78092912	17.0	113.5	70	971
	MAMIE	78100106	22.4	152.3	45	989
	ORA	78101112	20.9	127.4	55	980
	PHYLLIS	78101812	20.9	155.3	90	950
	RITA	78102200	11.2	153.5	115	946
	VIOLA	78111918	13.5	137.3	60	981

*Used in developing regression equations but not in testing because it had no official forecast error data.

INDEPENDENT SAMPLE

BESS	79032012	10.8N	139.9E	36	997
CECIL	79041406	8.4	130.4	63	983
DOT	79051118	10.2	122.3	26	1000
ELLIS	79070100	13.4	132.0	35	996
FAYE	79070400	10.3	139.9	46	991
HOPE	79072806	17.0	137.4	35	996
IRVING	79081406	24.0	124.9	80	956
JUDY	79081818	17.5	136.6	107	939
MAC	79091800	13.9	123.8	50	988
OWEN	79092500	20.9	131.0	72	965
SARAH	79100812	11.0	119.0	58	984
TIP	79100900	12.0	147.7	52	988
VERA	79110400	10.5	133.1	97	956
WAYNE	79110918	16.9	129.3	43	989
ABBY	79121006	11.8	133.1	64	979

would insure that the same number of analog storms are considered, and the similarity index weighting would be applied to both sets.

This limited comparison of the regression-adjusted forecast vector error standard deviations with those of the official forecast does seem to represent an increase in skill of the regression adjustment scheme over the official forecast.

V. CONCLUSIONS

The objective of this research was to adjust statistically the analogs selected by TYAN78 and produce improved forecast vector errors. Frill (1979) showed improvement in forecasts from the TCM with equations containing predictors based on backward integration of the TCM. A similar approach was used here in that the majority of the predictors selected in the stepwise regression program involved past position and motion information.

Multiple linear regression equations derived here reduce the mean and variance of the vector errors of tropical cyclone track forecasts. The inclusion of past position predictors in all of the equations indicates the role that persistence has in the predictions. However, the difference in the past 12-h motion of the storm and the analog storm is not simply extrapolated in time, as is the case in the present TYAN78 scheme.

An earlier version of the regression equations was based on only 11 input storms with only 161 analogs. Many of the same predictors were selected in that sample, but the coefficients were different. This set did not produce stable regression coefficients, because an independent sample resulted in degraded adjustments to the analog tracks.

If the TYAN78 model had not selected the input storm as an analog of itself, the sample size could have been increased. Thus, more stable regression equations could have been developed.

As described before, the regression equations were developed using analogs selected under the ARQ option for all storms. It is suggested that multiple linear regression equations be developed for each of the other options (recurver and non-recurver). Another possibility of improvement in the regression equations would be to use the earlier forecast position as a predictor for following positions. That is, one could explore the use of regression equations with the 24-h forecast position or related velocity components as a predictor for the 48- and 72-h forecasts. Similarily, the 72-h forecast would have available information regarding the 24- and 48-h forecast position. This would assure consistency in the adjustments at the various intervals.

APPENDIX A

METHODS:

The following section describes the methods used in obtaining and processing the data for this research.

Program NWPAC in TYAN78 was modified to allow the model to select and print out the top 20 analog storms. Subroutine NEXRPT was altered to force the TYAN78 program to print out all potential analog storms read during the historical search.

The output from the computer jobs at FNOC was manually searched for the top 20 analog storms. These were separated from the listing, marked with the rank and the "best" position identified. All analog storms listed were actual best tracks with positions given every six hours.

Most of the data was entered as card images on the W.R. Church computer center time-sharing system, then later punched onto cards. Once the raw data were on cards a track adjustment program was run to translate the analog storms. The track adjustment program produced punched cards with the translated positions.

The translated data were used in a predictor/predictand computing program which stored the computed values on disk. The BMDP2R program [Dixon and Brown, 1977] was run 42 times using the data from disk. The resulting equations are for intermediate times as well as the times described in the main text. All the equations were inserted in the evaluation program with resulting output described in the main text.

REFERENCES

1. Annual Typhoon Reports, 1977: U.S. Fleet Weather Central/ Joint Typhoon Warning Center, Guam.

2. Annual Typhoon Reports, 1978: U.S. Fleet Weather Central/ Joint Typhoon Warning Center, Guam.

3. Annual Typhoon Reports, 1979: U.S. Fleet Weather Central/ Joint Typhoon Warning Center, Guam.

4. Atkinson, G.D., Major, 1971: Forecaster's Guide to Tropical Meteorology. Tech. Report 240, Air Weather Service, Scott AFB, Illinois, p. 9-22.

5. Dixon, W.J. and M.B. Brown, 1977: Biomedical Computer Programs P-Series, Los Angeles; University of California Press.

6. Elsberry, R.L. and D.R. Frill, 1980: Statistical Post-processing of Dynamical Tropical Cyclone Model Track Forecasts, Mon. Wea. Rev.

7. Frill, D.R., 1979: Statistical Adjustment of Dynamical Tropical Cyclone Model Track Predictions, M.S. Thesis, Naval Postgraduate School, Monterey, California, 90 pp.

8. Jarrell, J.D. and R.A. Wagoner, 1973: The 1972 Typhoon Analog Program (TYFOON-72), Tech Paper 1-73, Naval Environmental Prediction Research Facility, Monterey, California 38 pp.

9. Ocean Data Systems, Inc., 1978: Final Report: Analog Tropical Cyclone Forecasting Program (TYAN78), 42 pp.

10. Ocean Data Systems, Inc., 1978: Program Maintenance Manual: Analog Tropical Cyclone Forecasting Program (TYAN78), 208 pp.

INITIAL DISTRIBUTION LIST

		No. of Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Dr. G.J. Haltiner, Code 63HA Chairman, Department of Meteorology Naval Postgraduate School Monterey, California 93940	1
4.	Air Weather Service Technical Library Scott AFB, Illinois 62225	2
5.	Captain William J. Kaveney AFIT/CIRF Wright-Patterson AFB, Ohio 45433	1
6.	Dr. R.L. Elsberry, Code 63Es Department of Meteorology Naval Postgraduate School Monterey, California 93940	8
7.	Captain M. Scott Risch Det 1, 1st WWg COMNAVMARIANAS Box 17 FPO San Francisco, California 96630	1
8.	Captain Kenneth A. Peterson 2d Weather Squadron/DA (CAPT Peterson) Andrews AFB, Maryland 20334	5
9.	Commanding Officer Naval Environmental Prediction Research Facility Monterey, California 93940	1
10.	Commanding Officer Fleet Numerical Oceanography Center Monterey, California 93940	1

11. CAPT Dennis R. Frill 1
Box 1946
APO New York, New York 09403

12. Commanding Officer 1
Joint Typhoon Warning Center
COMNAVMARIANAS Box 17
FPO San Francisco, California 96630

13. Ocean Data Systems, Inc. 1
2400 Garden Road
Monterey, California 93940

LME